A short note on the use of Mann Kendall statistics for detecting trends in atmospheric data

Sverre Solberg, Nilu, April 2015

Introduction

For the EMEP TFMM 2015 trend assessment the Mann Kendall (MK) trend methodology will be applied for estimating long-term trends in observational data. The plan is to apply this method to three time periods requiring a minimum data capture of 75%:

- 1990-2012
- 1990-2001
- 2002-2012

It is thus of interest to understand how appropriate the MK statistics are for our purpose. Basic questions to ask are:

- To what extent will the method detect trends for the given time periods?
- How reliable are the estimated Sen slopes?
- What is the effect of missing data?

To look into these questions, the MK methodoogy was investigated by use of artificial time series, corresponding in nature and length to the EMEP TFMM data. The artificial time series were constructed from two terms reflecting: i) A linear long-term trend due to changes in emissions and hemispheric baseline levels; and ii) A “natural variability” due to variations in meteorology from year to year. In reality, these two quantities could not be separated completely, but they express two different processes leading to variability in atmospheric observational data. Conceptually, it is clear that the magnitude of the long-term trend relative to the meteorological induced year-to-year variation is important for the success of detecting a trend with the MK method. The smaller the trend is relative to the inter-annual meteorological variability, the longer time series is needed in order to identify a significant trend. These relationships were studied below.

Method

Artificial time series were constructed based on the equation

$$C(y) = C_0 + fC_0\mathrm{RAND} - C_0\, a\, (y - y_0) \qquad (1)$$

$C(y)$	= Annual value at year y
C_0	= Annual value at the start year ($y = y_0$)
f	= amplitude (expresses the annual meteorological variability)
RAND	= rand number drawn from a standard normal distribution
a	= annual linear trend

Eq. 1 where then calculated with three settings of f and a, respectively, as given below:

f = [0.05, 0.10, 0.15]

a = [0.01, 0.02, 0.03]

This implies time series with an annual meteorological variability of 5, 10 and 15 % of the mean and a linear downward trend of 1, 2 and 3% of the initial concentration per year, respectively. Figure 1 shows examples of such artifical time series for a period of 13 years.

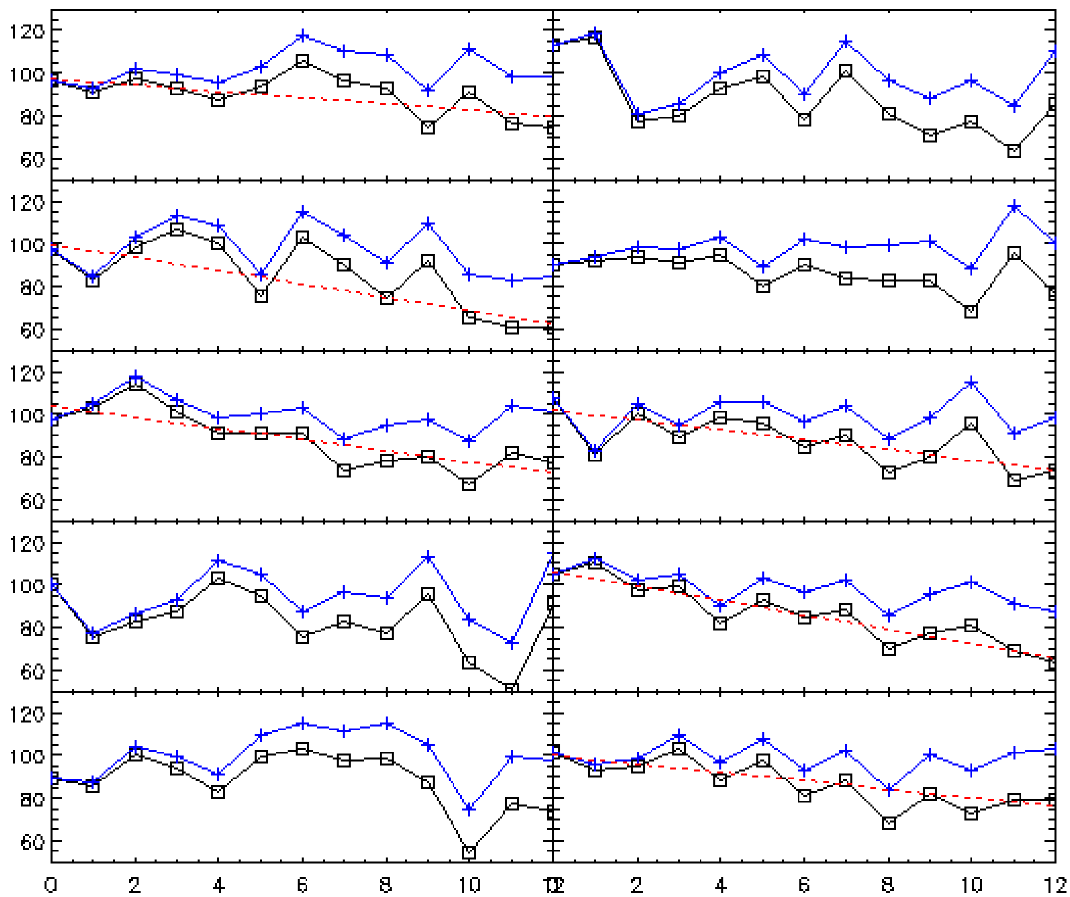

Figure 1. Examples of 10 artificial time series of annual values calculated for a period of 13 years with a linear trend of 2 %/year (relative to the start year) and a "natural variability" of 10 % st.dev. relative to the overall mean. The black line shows the time series with trend and natural variability, the blue line shows the time series without the trend and the red line gives the Sen's slope when the Mann Kendall methodology detects a significant trend (p=0.05).

The chosen values of the meteorological variability (5, 10 and 15 %) were based on EMEP model runs for the period 2000-2012 with fixed emissions (i.e. no emission changes from one year to another) and also on EMEP measurement data for various species.

From the model results, annual mean concentrations were calculated for each year and the st.dev. of these annual means divided by the overall mean were computed for each 50 x 50 km2 grid cell:

S = st.dev(Xi)/mean(Xi), Xi = annual mean in year i, i = [2000, …, 2012]

The range of S-values for the grid cells covering Europe were then calculated.

Similarly, S-values were estimated from annual mean observed concentrations of various species (SO_2, NO_2, PM10) during 2000-2012. Although these variabilities also include any long-term trends in addition to yearly fluctuations, we regarded the computed values as indicative of the range of the natural variablility.

Based on these modelled and observed data, we found that the range of 5 % - 15 % covered the typical range of the so-called natural variability with 10 % as a representative value for the median with some systematic differences between the various compunds. Lower variabilities were found for NO2 and higher variabilities for PM10.

To estimate the performance of the MK statistics, a number of N = 1000 random artificial time series for each combination of *f* and *a* were generated (Eq. 1), and then subject to a Mann Kendall test and Sen slope calculation. The calculations were done for time series with 11 and 23 years of data, respectively, corresponding to the time periods selected by the TFMM, i.e. 2002-2012 and 1990-2012. The fraction of these 1000 time series identified to have a significant trend (p=0.05) by the MK methodology and the mean of the corresponding slopes (based on the time series with a significant trend only) are given in Table 1 and Table 2. Table 1 gives the results with no missing data whereas Table 2 gives the results when assuming a maximum number of missing years (25%) of data, i.e. two and five missing years for the 11-years and 23-years time series, respectively. The missing years were distributed randomly in the time series.

Results

As seen from these results, the chances that the MK methodology detects a significant trend in a time series with 11 years of data are very small when the actual trend is of the order of 1 %/y. Furthermore, with a trend of 2 %/y and even 3 %/y one could only be fairly certain to find a trend when the natural (inter-annual) variability is small (5 %). With the strongest trend and a modest natural variability of 10 % it's likely (> 50 %) that the MK method will see the trend but the chances are not very high. The existence of missing years in the time series don't influence these results much. Assuming the maximum allowed amount of missing data only leads to a slight reduction in the MK performance.

The results reveal another important fact: In the cases of poor chances for the MK method to identify a trend (small trend and high natural variability), the absolute values of the estimated Sen slopes are consistently over-estimated. This may reflect that we in this exercise calculated the mean slope based on the cases with significant MK trends only, but it is nevertheless worth keeping in mind. The message seems to be that for time series of species/sites with uncertain trend estimations (a long-term trend being masked by the inter-annual variability), the estimated Sen slopes tend to be overestimated in absolute value.

With the longer time series of 23 years, the performance of the MK method was substantially improved. Furthermore, the existence of missing years had no influence on the results. With 23 years of data the chances that the MK method detects the trend is very good, except for the most difficult case (highest natural variability and smallest trend). The mean of the estimated Sen slopes are also very close to the real slope.

It is important to keep in mind that these results reflect very simplified conditions with a constant linear trend through the whole period and a random distribution of the natural variability. Real time series of atmospheric constituents will unfortunately not be like that. First of all, a constant (emission-induced) linear trend is not likely to to be seen, at least not over a period of 23 years. Secondly, systematic meteorology-driven inter-annual variations could complicate the trend analyses. As an example for ozone, the extreme year 2003 could bias the results substantially for certain choices of trend time periods.

St.dev	0.05		0.10		0.15	
trend	Sig. frac %	slope	Sig. frac %	slope	Sig. frac %	slope
-1	36	-1.4	13	-2.4	9	-3.0
	100	-1.0	81	-1.1	49	-1.4
-2	92	-2.1	37	-2.8	19	-3.7
	100	-2.0	100	-2.0	97	-2.1
-3	100	-3.0	71	-3.4	39	-4.1
	100	-3.0	100	-3.0	100	-3.0

Table 1. Results from Mann Kendall statistics on artifical time series with an annual trend of -1, -2, and -3 %/year and a natural variability corresponding to a st.dev of 0.05, 0.10 and 0.15 of the mean (see text for explanation). The numbers in the cells give the probability (Sig. frac %) that the Mann Kendall detects a significant trend (p=0.05) and the corresponding average Sen's slope (%/year). For each combination of trend and st.dev the calculations were done for 1000 artificial time series with 11 and 23 years of data, respectively (two lines from top to bottom in the cells). The combinations with a probability > 75 % for finding a trend is marked in green. These results are based on complete time series with no missing data.

St.dev	0.05		0.10		0.15	
trend	Sig. frac %	slope	Sig. frac %	slope	Sig. frac %	slope
-1	29	-1.4	12	-2.3	7	-3.1
	100	-1.0	90	-1.0	75	-1.1
-2	81	-2.1	33	-2.8	16	-3.8
	100	-2.0	100	-2.0	98	-2.0
-3	98	-3.0	60	-3.4	31	-4.3
	100	-3.0	100	-3.0	100	-3.0

Table 2. Same as Table 1 but with 25 % missing years in each time series, i.e. two missing years in the 11-years data and five missing years in the 23-years data.

Conclusion

This small exercise has revealed that the length of the time series is critical for the performance of the MK statistics. With time series of annual values and only 11 years of data, the chances that the MK method detects the long-term trend is small, except for the cases with a small annual variability and a marked trend (3 %/y or stronger in absolute level). It is furthermore a risk that the absolute value of the calculated trend is overestimated. With 23 years of data, the situation is very different, and the MK method will most certainly detect the trend, and the estimated slopes will be close to the real trend in the data.